

~~CONFIDENTIAL~~

24 August 1960

MEMORANDUM FOR: Assistant Director, Central Reference

SUBJECT : Intellofax System

I. PROBLEM

To determine organizational requirements of a modified Intellofax System.

II. FACTS BEARING ON THE PROBLEM

1. The revised edition of the ISC, together with the alternative Area Classification Codes, has been approved by CODIB and distributed to the Intelligence Community.
2. The existing Intellofax file contains nine million cards. They cannot be merged effectively with cards coded under the new ISC.
3. The Air Force has been Minicoding according to the old ISC and expects to start using the revised codes in late 1960. There is every indication that their input will be compatible with our coding.
4. The Intellofax Task Group, composed of representatives from the Document and the Machine Divisions, the CIA Library and the Automation Development Group has submitted a study on modification of indexing requirements and techniques as well as expansion of the Intellofax files to provide for uniform and specific retrieval. The recommendations of the Intellofax Task Group are endorsed by the undersigned to the extent indicated in this paper. A copy of this study is attached as Tab A.
5. Only the manpower available to the three Divisions concerned has been considered.

III. DISCUSSION

1. We believe that OCR should lead the way by putting the revised subject and area codes into effect for the Intellofax System as soon as possible to provide customers with the most effective retrieval system. Adoption of these codes will provide an opportunity to take advantage of both Intellofax and Minicard coding experience and to make certain desirable modifications in input and retrieval procedures.

~~CONFIDENTIAL~~

CONFIDENTIAL

-2-

2. The preparation of a coding manual will be required before the new codes can be put into effect. This is essential for coding (indexing), and for processing retrieval requests. A draft of a coding manual is being drawn up by the Task Group; its scope will be dependent in part on the approval of this paper. At least thirty days will be required after completion of the draft to train document analysts and librarians in the use of the revised ISC and the Dictionary. The Machine Division will requisition a supply of the revised punched card stock as soon as the proposed card format is accepted. If the manufacturer fails to make delivery to meet our schedule, the new system can nevertheless be initiated (the reverse side of present stock of cards can be used). During the training period and immediately following adoption of new procedures, the Document and the Machine Divisions will require substantial amounts of overtime to prevent backlogs.
3. When indexing with the new codes is started, a new Intellofax file will be created and will have to be searched in addition to the present Intellofax file in processing retrieval requests. The simultaneous maintenance and use of two files is inherent in the decision to adopt the new ISC. It will be several years before traffic in the old file drops off significantly; by this time it is expected that increased use of the Intellofax System will more than offset this slight saving. Additional manpower in the Machine Division, amounting to approximately two persons, will be required to search the two files simultaneously.
4. There is general agreement among the three divisions concerned and the ADG that the four digit numeric area notation should be used in preference to the six digit alphabetic notation. The four digit numeric scheme is more economical machine-wise and since the two notations approved by CODIB are interchangeable item-for-item, all systems using one or the other will be compatible.
5. Subjects dealing with several areas pose a special problem. In the present subject file primary and secondary areas are recorded only on one card except in special cases. A search for secondary areas requires more machine sorting time and retrieval scanning by the Composite Group. The new area code and proposed card format permit automatic reversal of all areas. To accomplish this equal area status in the subject file would increase its size by approximately twenty-five per cent and would require approximately one half additional person in the Machine Division.

CONFIDENTIAL

CONFIDENTIAL

-3-

6. The use of a dictionary is essential in obtaining greater uniformity in coding input and results in better retrieval. The usage decisions cumulated during the Minicard test are being used as the basis for the dictionary. It is expected that the dictionary will grow steadily for more than a year and the maintenance will always require increment of manpower. The review of indexing performed by Document Division analysts will be accomplished by strict adherence to the dictionary procedures and by review of analysis as deemed necessary by the Document Division. For this, Document Division will need three persons and Machine Division one person on a continuing basis.
7. An area file which completely duplicates the subject file would add significantly to the reference capability of OCR. It would permit half hour service instead of two day service on perhaps 30 per cent of current requests, and would encourage spot requests which are not directed to the Intellofax System at present. Chief, Machine Division, estimates that he would require 3 additional personnel for this purpose. Estimates of need and suitability of this type of service, as well as the operating efficiency of maintaining area and subject files simultaneously can be more meaningfully made after a test period. Such a test can be made at no additional cost, by arranging in area order the VMR set of cards prepared under the new ISC. A final decision on the establishment of an area file should be postponed until such a test is made.
8. In searching by subject, many codes in the new ISC will lead to large numbers of documents pertinent to the same general subject. As a result, review of a great number of card references is required to identify pertinent references. Specificity among these reports can in many cases be defined and can be brought under machine control. The use of clear text for this purpose, controlled by a dictionary of terms actually used, would assure a high degree of uniformity, yet without the need for frequent revision of the basic code book. The use of a dictionary controlled clear text on a highly selective basis for machine selection purposes would require more coding and dictionary time as well as more key punching and card filing during the input phase. The extent of its use as set forth below is based on available manpower and retrieval experience. Expansion of titles and preparation of abstracts for the benefit of consumers in selecting documents from Intellofax tapes are not affected by the use of the punched clear text. The following uses of clear text on a highly selective

CONFIDENTIAL

CONFIDENTIAL

-4-

and controlled basis as determined by the three divisions concerned are recommended:

- a. Subject definition: To identify subclasses included in a single ISC code entry. For example "coking" coal within coal; "admiralty law" within International Law.
- b. Corporate names: Political parties, organizations, and conferences only. Installations of the categories covered by the Industrial Register will not be listed in clear text.
- c. Geographic place names: Coverage controlled by the Dictionary List Out and the Coding Manual.
- d. Personal names: The Intellofax will not substitute for the biographic approach afforded by BR and BI/State. Therefore, the tag for persons will be used only to indicate the presence of names in a document.

The net manpower cost of use of clear text is especially difficult to estimate. It will be several months before criteria are established for inclusion or exclusion of types of terms and for establishment of authoritative forms of entry for the terms used. Our best guess is that one person would be required in the Machine Division, and one in the Document Division.

9. The proposed IBM card form will accept and permit machine control of specific information in detail. The proposed card accepts all entries presently in use, so that its adoption will impose no limitation on current capacity. The proposed format for the code sheets and punch card is designed to facilitate input to the system.
10. The Air Force plans to adopt the new codes, but has not yet completed their revised code manual and procedures and the date of change over has not been determined. Ability to continue to convert Air Force Minicoding for Intellofax input is vital to manpower savings in the Document Division. Close liaison is being maintained with Air Force to insure coding compatibility. The machine conversion of minicoding will be less complex with our proposed system. The extent of simplification will be determined by the course that Air Force takes. The Air Force is also planning a dictionary and the Document Division plans to develop procedures whereby index cards can be exchanged and listings be supplied to the Air Force to obviate needless duplication of work.

CONFIDENTIAL

CONFIDENTIAL

-5-

11. The gross cost of implementing all of the recommended changes without regard to possible economies through realignment of certain functions, would be approximately thirteen additional persons. The CIA Library would require the equivalent of one person to code for retrieval from both the old and the new files and would absorb this requirement in its present T/O. The Document Division would save two persons by reorganizing the Analysis Branch and another by ceasing to code those finished intelligence series that are included in the IPI. Staff previously assigned to Minicard (5-Machine, 1-CIA Library) would be made available for the proposed program. The net increase in T/O in the three operating Divisions concerned would therefore be three persons. (Tab B.)
12. Full-time representation by the Document Division and consultive representation by the Machine Division on the Composite Group provide three way communication that does much to improve quality of input and of output. This should by all means be continued. Many input decisions are now being formalized in the coding manual with mutual participation by the Document Division, by the Machine Division, and by the Library. These three divisions should continue to be parties to major input decisions affecting the speed and quality of Intellofax retrieval such as criteria for the use of clear text, or the information to be included on the face of Intellofax cards, or the "nodex" criteria.

IV. RECOMMENDATIONS

1. That the numeric area notation be used with the revised ISC.
2. That the coding manual and the dictionary be adopted and maintained as standard tools for document processing and retrieval.
3. That dictionary controlled clear text be used to the extent outlined in this paper until such time that an extension is deemed necessary by the three divisions concerned.
4. That review of indexing input be performed on a regular basis to the extent considered necessary by the Document Division.
5. That the indexing of finished intelligence be eliminated from the Intellofax System and be serviced through reference to the IPI.
6. That a card be filed in the subject file for each area code whether it be "primary" or "secondary."

CONFIDENTIAL

CONFIDENTIAL

-6-

7. That the three divisions (Document, Machine, and Library) continue to participate in major input decisions affecting retrievability.
8. That Machine Division T/O be increased by three, the CIA Library T/O be decreased by one which is to be transferred to the Document Division to offset the assignment of a document analyst to the Library Composite Group.
9. That the revised Intellofax System be put into operation within OCR as soon as possible.

25X1

*Believe 1 October 1960
should be announced as
chargeover date. DD has
already started training and
considers that date realistic.
Pult*



Attachments: (2 Tabs)
As stated above.

25X1

Concur:

Approv



Aug. 30, 1960.

CONFIDENTIAL